PNAS
www.pnas.org

Supplementary Information for

Hippocampal CA1 gamma power predicts the precision of spatial memory judgments

Rebecca F. Stevenson[1,2], Jie Zheng[3], Lilit Mnatsakanyan[4], Sumeet Vadera[5], Robert T. Knight[6,7], Jack J. Lin[2,3,4], Michael A. Yassa[1,2,4]

[1]Department of Neurobiology and Behavior, University of California, Irvine
[2]Center for the Neurobiology of Learning and Memory, University of California, Irvine
[3]Department of Biomedical Engineering, University of California, Irvine
[4]Department of Neurology, University of California, Irvine
[5]Department of Neurological Surgery, University of California, Irvine
[6]Department of Psychology, University of California, Berkeley
[7]Helen Wills Neuroscience Institute, University of California, Berkeley

Michael A. Yassa
Email: myassa@uci.edu

Jack J. Lin
Email: linjj@uci.edu

This PDF file includes:

Supplementary text
Figs. S1 to S8
References for SI citations

www.pnas.org/cgi/doi/10.1073/pnas.1805724115

Supplementary Information Text

Materials and Methods

Spatial precision task

One hundred images of common objects were selected from a set previously used by our group (1). Images were presented on a laptop computer screen set at a comfortable distance from the patient. Three hundred and sixty locations on the computer screen were generated along the circumference of a circle centered on the screen with a spacing of 1°. During encoding, the 100 objects appeared one at a time at pseudorandomly assigned circle locations and subjects were asked to judge whether each object would more likely be found indoors or outdoors (Figure 1; 1.2 +/- 0.2 second ITI; .5 second ISI). Following a short (~1 minute) delay, subjects were shown each object again in pseudorandom order, this time at the top of the screen. Subjects were instructed to wait 1 second (until text that read 'Wait…' disappeared from the screen) before using a mouse wheel to move the object to where it appeared during encoding. Subjects pressed the space bar to indicate that they were finished placing the object. If subjects performed more than one session, new objects and locations were used. No more than one session was performed on each day.

Behavioral analysis

Error on the spatial precision task was measured as the number of degrees between where subjects placed the object and the correct location. Histograms were used to examine the distribution of error values. We used mixture modeling, as implemented by the MemFit function of Memtoolbox (2), to obtain an estimate of two parameters describing these distributions: the guess rate (g), which reflects the area under the uniform distribution, and the standard deviation of the von Mises distribution (SDMem). We used the cumulative distribution function of the von Mises distribution estimated for each session to split trials into three conditions: High precision, Low precision, and Guess. Trials that had less than a 10% chance of being remembered with some degree of precision were placed in the Guess condition. The remaining trials were sorted by error and split evenly into the High and Low precision conditions.

Electrode localization

The electrode localization was performed using pre- and post-implantation structural T1-weighted 1mm isotropic MRI scans as well as post-implantation CT scans. For each participant, the post-implantation MRI and CT scans were registered to the pre-implantation scan using a 6-parameter rigid body transformation implemented with Advanced Normalization Tools – ANTs (3). Electrodes were localized within MTL subregions using a high-resolution (.55 mm) in-house anatomical template with manual tracings of hippocampal subfields and parahippocampal gyrus subregions (4). Regions of interest (ROIs) included the CA1, DG/CA3, subiculum (Sub), lateral and medial entorhinal cortex (LEC, MEC), and the perirhinal (PRC) and parahippocampal (PHC) cortices. Hippocampal subfield segmentation followed our previously published protocols (4). The labeled template was resampled and aligned to each subject's pre-implantation scan using ANTs Symmetric Normalization, so that the labels could be used to guide localization. Each electrode location was determined by examining the co-registered pre- and post-implantation MRIs and identifying the ROI that corresponded to the center of the electrode artifact in the post-implantation MRI and CT. Cases in which electrodes were on the border between ROIs or between gray matter and white matter were noted as such. Outside the MTL, electrode localization was guided by a FreeSurfer cortical parcellation of the pre-implantation MRI (5).

Data collection and preprocessing

Intracranial EEG data were recorded using a Nihon Khoden recording system, analog-filtered above 0.01 Hz and digitally sampled at 5000 Hz. After acquisition, data were demeaned and band-pass filtered from 0.3 Hz to 350 Hz using a two pass zero phase delay Butterworth infinite impulse response (IIRR) filter. Power spectra were examined to identify line noise and a Butterworth notch filter was used to remove 60 Hz noise and harmonics. All electrodes were re-referenced to a white matter electrode located on the same depth electrode probe. A neurologist (J.L.) with subspecialty training in epilepsy visually inspected continuous recordings from each session to identify all data with interictal epileptiform discharges. Data were also inspected for excessive noise, including broadband electromagnetic noise from hospital equipment. To avoid potentially biasing the results, the neurologist was blinded to trial information (e.g. stimulus onset and behavioral performance) as well as to electrode location. Only data from recordings contralateral to the seizure source or outside of the seizure onset zone were used in subsequent analyses.

Gamma power analyses

Intracranial recordings were broken into event-related epochs (3 seconds pre-stimulus onset and 3 seconds post-stimulus offset) and convolved with complex Morlet wavelets, implemented using the FieldTrip toolbox, to obtain a measure of instantaneous power (6). Center frequencies ranged from 1 to 150 Hz, with a spacing of 1 Hz and a variable cycle number of 4-15. Power was baseline corrected to the average pre-stimulus power across all trials (0.5 to 0.2 prior to stimulus onset), resulting in a measure of relative power per frequency (power divided by pre-stimulus power) at each time point. While we examined a wider range of power in the spectrograms (up to 150 Hz), we used an a priori gamma frequency range of 40-100 Hz for our gamma power analyses. This frequency range was based on prior literature showing MTL gamma activity in this range (7, 8, 9) and reflected the dominant gamma frequency range that was influenced by task performance (Fig. S2). For the within-session analysis, we then averaged baseline corrected power over our gamma frequency range (40-100 Hz) and retrieval window (0.25-1 second post-stimulus onset). We took the absolute value of the angular error and logged the resulting values to account for the non-normal distribution of error. Pearson correlation was used to test the relationship between single trial power in each MTL electrode and error. P values from these correlations were then Bonferroni-corrected for the number of MTL electrodes in each patient (see Table 1 for the number of electrodes). For the across-session analysis, electrodes were divided into nine regions: 1) the CA1 subfield, 2) the hippocampus, including the CA1, 3) the entorhinal, perirhinal, and parahippocampal cortices, 4) the lateral temporal cortex, 5) the insula, 6) the caudal prefrontal cortex (broadman area (BA) 6/8), 7) the orbitofrontal cortex, 8) the anterior cingulate cortex, and 9) the dorsolateral prefrontal cortex (BA 9/10/46). Electrodes that were on the border between the CA1 subfield and other hippocampal subfields were included in the hippocampal region but not in the CA1 region. Regions where we had electrode contacts in fewer than three subjects (e.g. amygdala) were excluded from the analysis. There were not enough electrodes across subjects in either the DG/CA3 or subiculum to be able to look at activity separately within these regions. Power was baseline corrected to the average pre-stimulus power (as described above), and subsequently z-transformed separately within each session to account for differences in power and noise across sessions.

We used a cluster-based permutation approach implemented using the FieldTrip toolbox to examine the correlation between gamma power and error (including the Guess condition) and precision (excluding the Guess condition) at each time point within each region (6, 10). We also

used this method to examine the relationship between theta (3-8 Hz) power and error in each region and found a significant negative correlation between theta power and error during the retrieval window in CA1 and a significant positive correlation that emerged after the retrieval window in dlPFC (Fig. S4). Neither of these regions showed a significant correlation between theta power and precision. For Figures 5d-f and 6b, trials were divided into High precision, Low precision, and Guess conditions, then z-transformed power was averaged over the retrieval time window (0.25- 1s post-stimulus onset) and a one-way ANOVA was performed across the three conditions. In regions where we found a main effect of error, we tested for pairwise differences between conditions using post hoc Holm-Sidak tests.

The number of expected guesses in the High and Low precision conditions (Fig. S5) was balanced using the following method. The number of degrees spanned by each condition was calculated for each session and used to determine the number of expected guesses in each condition as predicted by the uniform distribution. If there were more expected guesses in the Low precision condition, we added simulated 'guesses' to the High precision condition so that the number of High and Low precision guesses would be equal. These simulated 'guesses' were generated by taking the mean power in the Guess condition during the retrieval window. We also added the same number of simulated 'null' trials to the Low precision condition to keep the trial count balanced across conditions. These 'null' trials were generated by taking the mean power in the Low precision condition during the retrieval window.

Granger Prediction Analysis

The preprocessed LFP was first downsampled to 250 Hz before obtaining the mean gamma power time series (40-100 Hz) during the full retrieval period (0-1s post stimulus onset). These data were epoched, detrended, and normalized over time and across trials to increase stationarity (11). A model order of five was chosen for each session. This order was determined by the Bayesian information criteria, which was calculated using the Multivariate Granger Causality (MGVC) Matlab Toolbox (12). The MGVC toolbox was also used to calculate the time-domain Granger prediction index for High precision, Low precision, and Guess trials for each CA1 and dlPFC electrode pair. The difference in Granger prediction values between conditions (e.g. High-Guess) was calculated for each electrode pair, averaged over electrode pairs within each session for each direction (CA1 to dlPFC and dlPFC to CA1), and then averaged across sessions. A null average difference distribution was created by shuffling the trial labels 500 times before calculating the difference in Granger prediction values between conditions. These distributions of permuted difference values were then averaged first over electrode pairs and then across sessions, as described above (13). The observed average Granger difference value was compared to this null average Granger difference distribution. P-values were calculated as the fraction of times the null average Granger difference values were equal to or more extreme than the observed average Granger difference value.

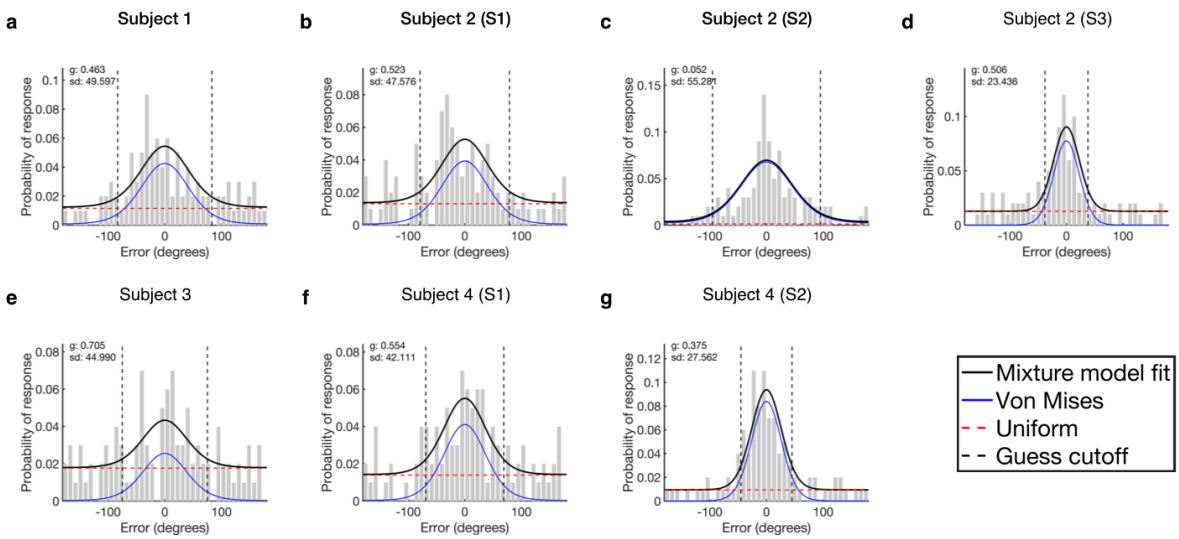

Fig. S1. Error histograms, mixture model fit, and cutoff values for the Guess condition for each session. a-g, The black curve indicates the mixture model fit, the blue curve indicates the pdf of the Von Mises distribution, the red dashed line indicates the pdf of the uniform distribution, and the black dashed lines indicate the cutoff values for the Guess condition. g: guess rate; sd: standard deviation of the Von Mises distribution (SDMem); pdf: probability density function; S1: Session 1; S2: Session 2; S3: Session 3.

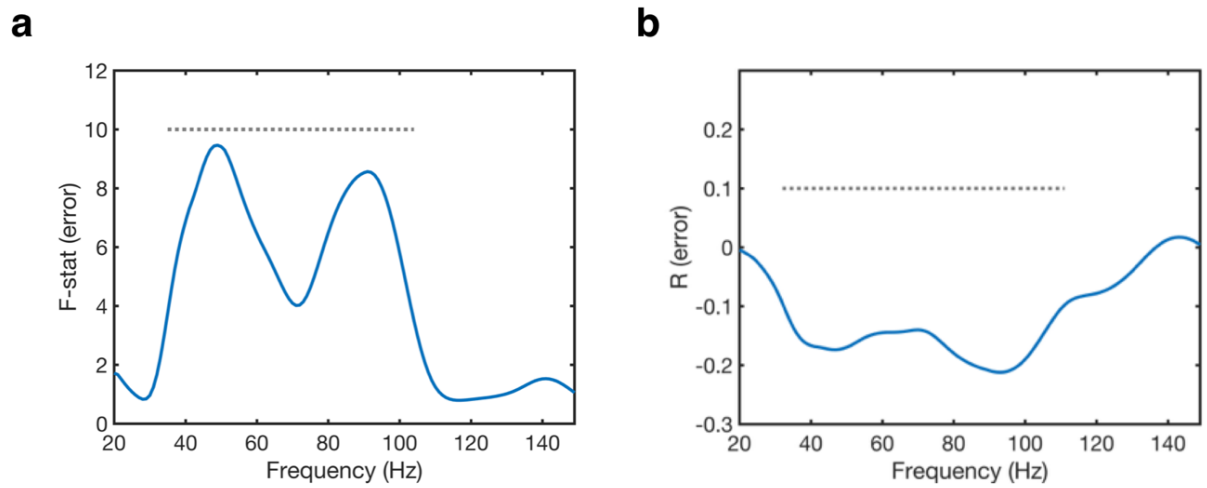

Fig. S2. Gamma frequency range and task performance. We calculated the mean CA1 power during the retrieval window at successive gamma frequency ranges using a sliding window moving average of 20 Hz. The center frequencies showing a significant effect of error across conditions (High precision, Low precision, and Guess) ranged from 35 to 104 Hz, consistent with our a priori selection of 40-100 Hz as our gamma frequency range of interest **(a)**. The center frequencies showing a significant correlation between gamma power and error **(b)** covered a similar range (32 to 111 Hz). Dotted gray horizontal line indicates $p < 0.05$.

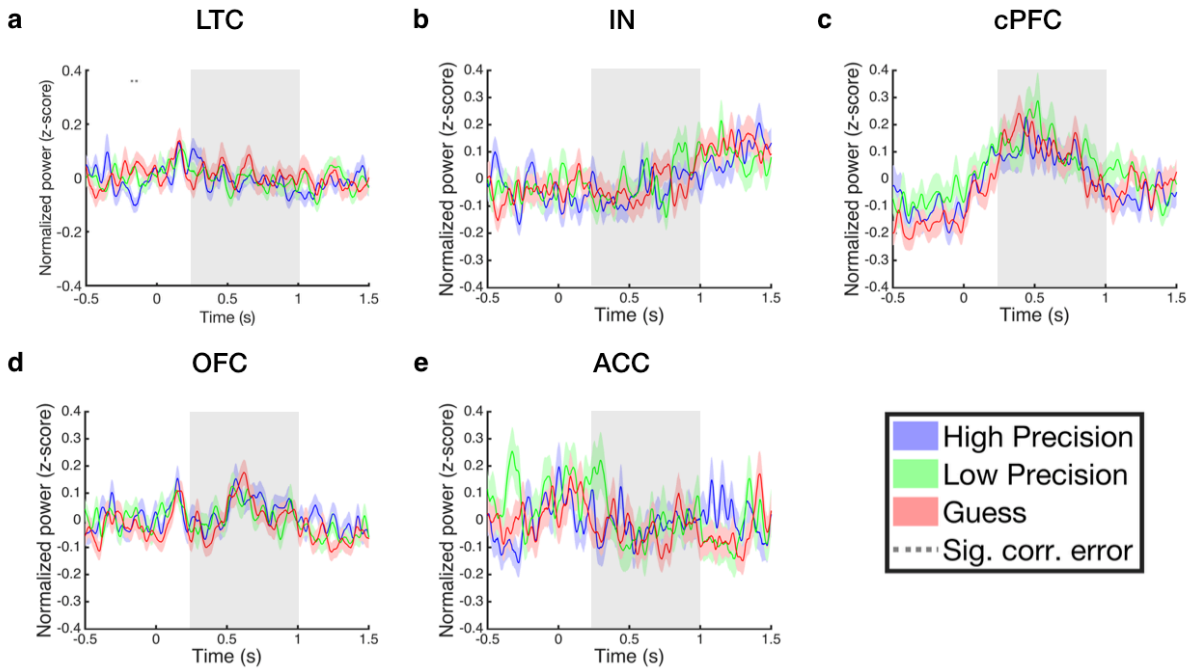

Fig. S3. **Gamma power and error.** Time course of gamma power in the lateral temporal cortex **(a)**, insula **(b)**, caudal prefrontal cortex (Brodmann areas 6/8) **(c)**, orbitofrontal cortex **(d)**, and anterior cingulate cortex **(e)**. Stimulus onset is at time zero and the retrieval window (0.25 to 1 second post-stimulus onset) is shaded in gray. Gray horizontal lines indicate time points where there are significant correlations between gamma power and error ($p < 0.05$). The lateral temporal cortex showed a significant effect of error pre-stimulus onset. Colored shaded regions indicate s.e.m.

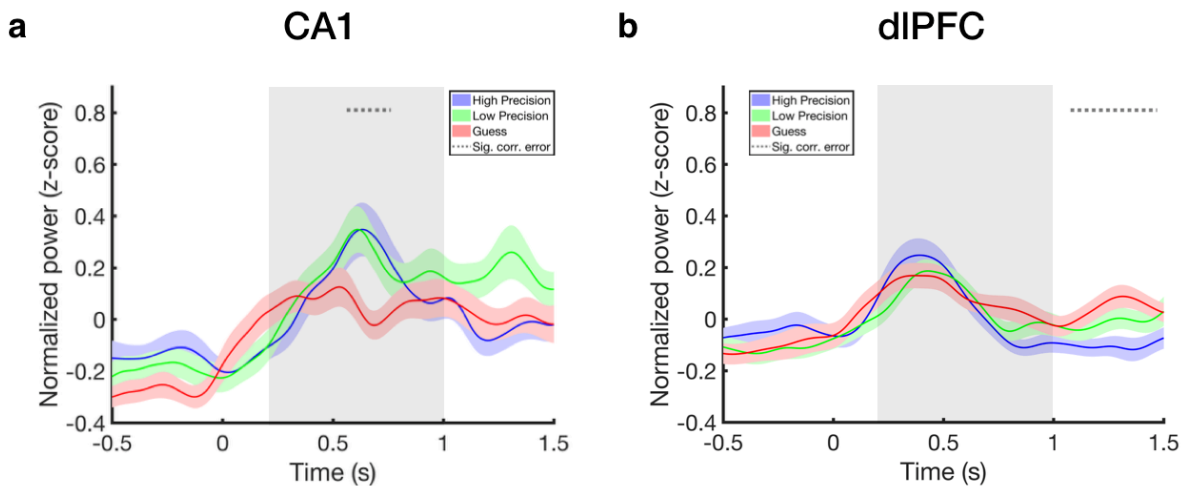

Fig. S4. Theta power and error. Time course of theta (3-8 Hz) power in the CA1 subfield **(a)** and dlPFC **(b)**. Stimulus onset is at time zero and the retrieval window (0.25 to 1 second post-stimulus onset) is shaded in gray. Dotted gray horizontal lines indicate time points where there are significant correlations between gamma power and error ($p < 0.05$, cluster-corrected). Colored shaded regions indicate s.e.m. We found a significant negative correlation between theta power and error during the retrieval window in CA1 and a significant positive correlation that emerged after the retrieval window in dlPFC. Neither of these regions showed a significant correlation between theta power and precision.

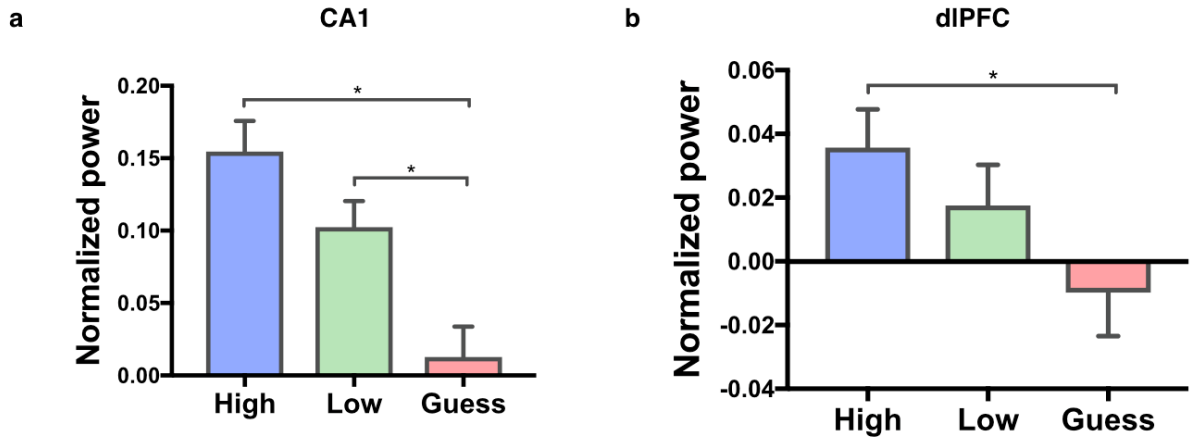

Fig. S5. Gamma power in the CA1 and dlPFC: number of expected guesses balanced across High and Low precision conditions. Mean gamma power over the retrieval window (.25 to 1 second post-stimulus onset) in the CA1 **(a)** and dlPFC **(b)** for High precision, Low precision and Guess conditions with the number of expected guesses balanced across High and Low precision conditions. We ran a one-way ANOVA across conditions (High, Low, and Guess) in each region. The results of these ANOVAs were similar to those obtained without balancing (CA1: $F(2, 462) = 12.5$, $p = 5 \times 10^{-6}$; High vs. Guess: $p = 3 \times 10^{-6}$; High vs. Low: $p = 0.06$; Low vs. Guess: $p = 0.004$; dlPFC: $F(2, 651) = 3.19$, $p = 0.04$; High vs. Guess: $p = 0.037$; High vs. Low: $p = 0.32$; Low vs. Guess: $p = 0.25$).

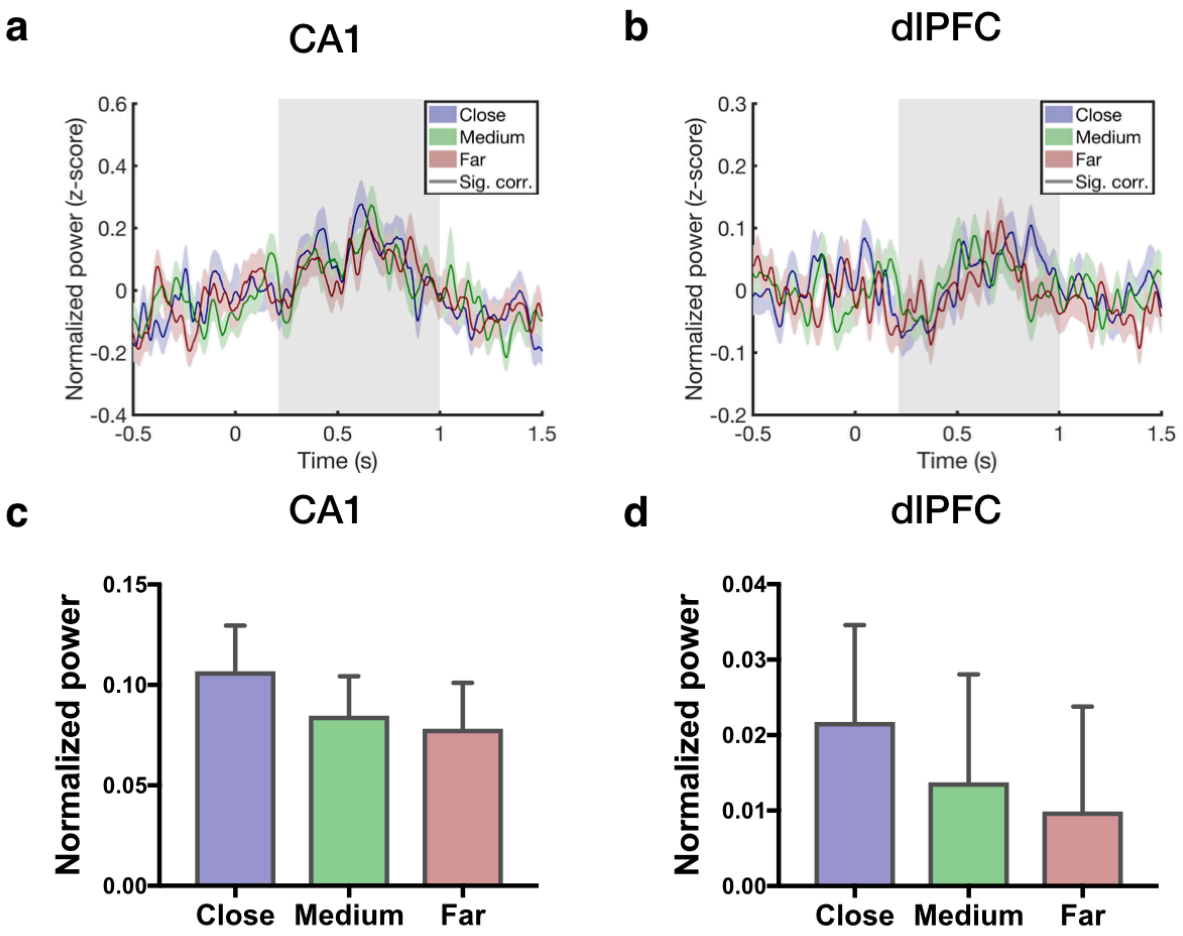

Fig. S6. Gamma power is unrelated to the distance object was moved. We examined the relationship between gamma power and the distance the object was moved by calculating the angular distance from the top of the screen (the starting point) to where the subjects placed each object. Trials were sorted by this distance and divided evenly into three conditions: Close, meaning objects that were placed closer to the top of the screen, Medium, meaning objects that were placed a medium distance from the starting point, and Far. The time course of gamma power is shown in the CA1 **(a)** and dlPFC **(b)** for Close, Medium, and Far Trials. No significant correlations between gamma power and distance were found at any time point ($p > 0.05$, cluster corrected). Additionally, CA1 **(c)** and dlPFC **(d)** gamma power was averaged over the retrieval window and a one-way ANOVA was performed across the three conditions. These ANOVAs did not show a significant effect of distance in either region (CA1: $F(2, 426) = 1.1$; $p = 0.32$; dlPFC: $F(2, 603) = 0.78$; $p = 0.46$). Colored shaded regions and error bars indicate s.e.m.

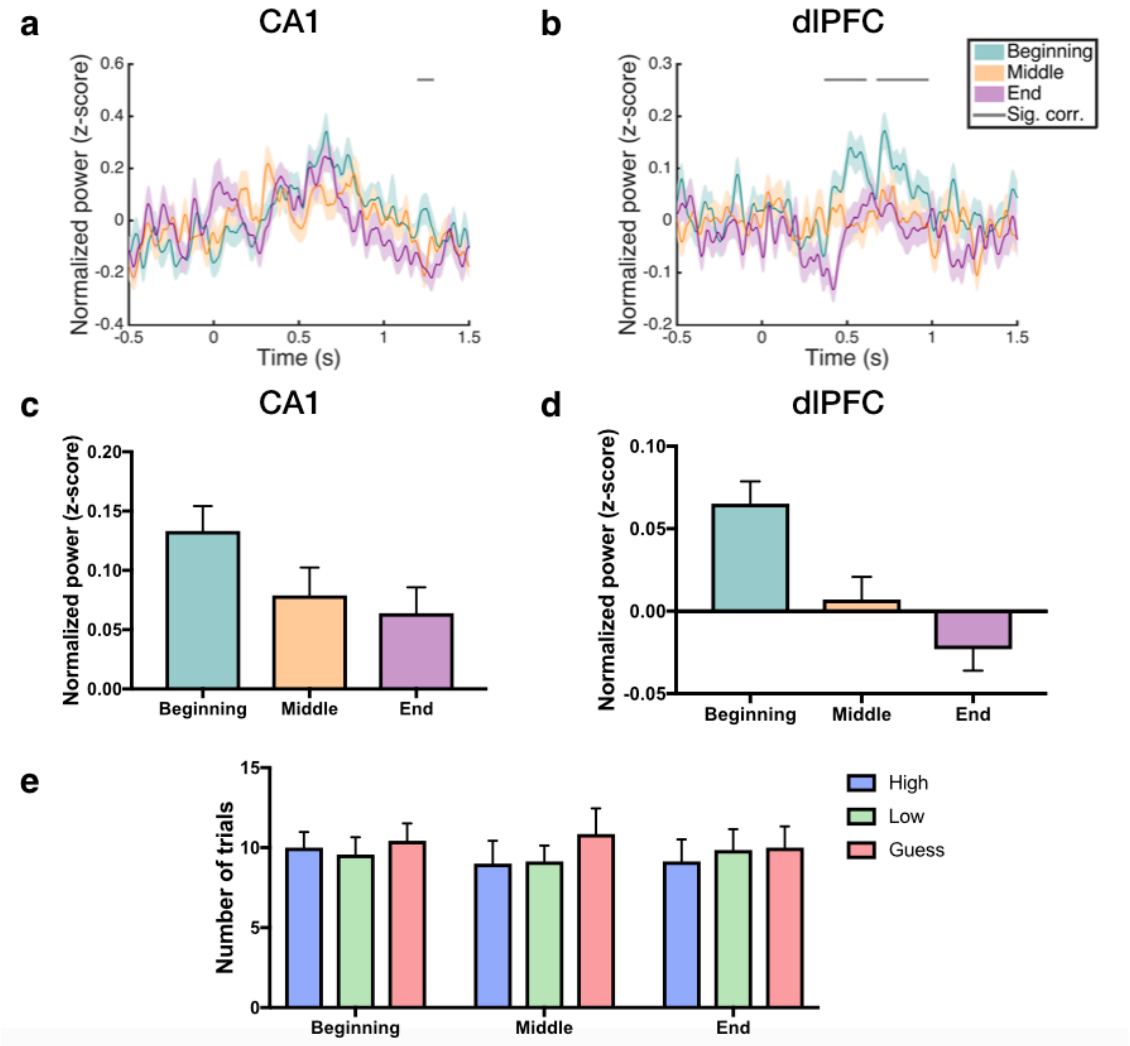

Fig. S7. Gamma power decreases over the course of the test session do not drive the error/precision effects. To assess the relationship between gamma power and trial order, trials were divided evenly into three conditions: Beginning, meaning trials from the first third of the test session, Middle, and End. The time course of gamma power in the CA1 **(a)** and dlPFC **(b)** is shown for trials at the Beginning, Middle, and End of the test session. Gray horizontal lines indicate time points where there are significant correlations between gamma power and trial order ($p < 0.05$, cluster corrected). Colored shaded regions indicate s.e.m. c, d, Mean gamma power over the retrieval window (.25 to 1 second post-stimulus onset) for trials at the Beginning, Middle, and End of the test session in the CA1**(c)** and dlPFC **(d)**. One-way ANOVAs with trial order (Beginning, Middle, and End) as fixed factors revealed a significant effect of order in the dlPFC ($F(2, 603) = 11.2$; $p = 0.00002$) and a marginal effect in the CA1 ($F(2, 426) = 2.7$; $p = 0.07$). Error bars indicate s.e.m. **e**, To assess the relationship between trial order and error, the number of High precision, Low precision, and Guess trials was calculated for each condition (Beginning, Middle, and End), and a two-way ANOVA was performed with trial order (Beginning, Middle, and End), and error (High, Low, and Guess) as fixed factors. This ANOVA resulted in p values > 0.05 (Trial order: $F(2, 12) = 0.5$, $p = 0.6$; Error: $F(2,12) = 0.3$, $p = 0.7$; Interaction: $F(4, 24) = 0.2$, $p = 0.9$), indicating that the number of High precision, Low precision, and Guess trials does not change over the course of the test session.

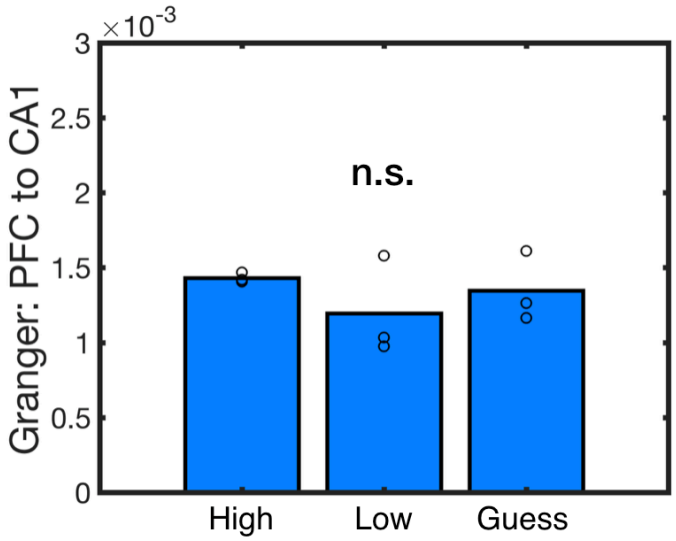

Fig. S8. Mean dlPFC to CA1 Granger prediction values for each condition. P > 0.05 as determined by permutation testing (see SI Materials and Methods).

References

1. Stark, S.M, Stevenson, R., Wu, C., Ruteledge, S., Stark CE (2015) Stability of age-related deficits in the Mnenomic Similarity Task across task variations. *Behav Neurosci* 129(3):257–268.
2. Suchow JW, Brady TF, Alvarez GA (2013) Modeling visual working memory with the MemToolbox. 13:1–8.
3. Avants BB, et al. (2011) A reproducible evaluation of ANTs similarity metric performance in brain image registration. *Neuroimage* 54(3):2033–2044.
4. Yassa MA, Stark CEL (2009) A quantitative evaluation of cross-participant registration techniques for MRI studies of the medial temporal lobe. *Neuroimage* 44(2):319–327.
5. Fischl B, et al. (2004) Automatically Parcellating the Human Cerebral Cortex. *Cereb Cortex* 14(1):11–22.
6. Oostenveld R, Fries P, Maris E, Schoffelen JM (2011) FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput Intell Neurosci* 2011. doi:10.1155/2011/156869.
7. Greenberg J a., Burke JF, Haque R, Kahana MJ, Zaghloul K a. (2015) Decreases in theta and increases in high frequency activity underlie associative memory encoding. *Neuroimage* 114:257–263.
8. Sederberg PBP, et al. (2007) Gamma oscillations distinguish true from false memories. *Psychol Sci* 18:927–32.
9. Burke JF, et al. (2014) Human intracranial high-frequency activity maps episodic memory formation in space and time. *Neuroimage* 85:834–843.
10. Maris E, Oostenveld R (2007) Nonparametric statistical testing of EEG- and MEG-data. *J Neurosci Methods* 164(1):177–190.
11. Seth AK (2010) A MATLAB toolbox for Granger causal connectivity analysis. *J Neurosci Methods* 186(2):262–273.
12. Barnett L, Seth AK (2014) The MVGC multivariate Granger causality toolbox: A new approach to Granger-causal inference. *J Neurosci Methods* 223:50–68.
13. Solomon EA, et al. (2017) Widespread theta synchrony and high-frequency desynchronization underlies enhanced cognition. *Nat Commun* 8(1). doi:10.1038/s41467-017-01763-2.